Multilingualism and language shift in a Malaysian Hakka family

Ralf Vollmann & Tek Wooi Soon

University of Graz, Institute of Linguistics; ralf.vollmann@uni-graz.at

Abstract.

Background. In multiethnic and multilingual Malaysia, four standardized languages (Malay, English, Chinese, Tamil) and a number of spoken languages (e.g. Hakka, Bahasa Pasar, Malaysian English) serve pluriglossic purposes. Today, Standard (Malaysian) Chinese is used by ethnic Chinese not only at the acrolectal, but also at the mesolectal (inter-group communication) and basilectal level (family, friends). Studies have observed language shift of the family language from a smaller Chinese language to Malaysian Mandarin.

Material & method. This study investigates the language use in one Hakka family living in Peninsular Malaysia (KL, Penang) and mainland China. Specific focus lies on Hakka as the family language and its intergenerational development.

Analysis. All family members are multilingual, the older and middle generations use various languages beside their family language, Hakka. The middle-aged speakers consider their own Hakka 'impure'; code-switching and multilingual conversations are a regular occurrence. Interethnic marriages and school education make Mandarin the most suitable language for communication even in the family. Thus, Hakka is no longer actively acquired by the youngest family members.

Conclusions. Hakka is viewed as the family language and still widely used, but its position is substantially weakened in the youngest generation.

Keywords. Hakka, Malaysia, minority language, multilingualism, language loss

The authors want to thank Dörte Borchers for comments on an earlier version of this paper.
1. Background

1.1. The linguistic situation in Malaysian Chinese

Malaysia is a multi-ethnic society with many languages (Asmah 1992; Vikør 2004). Malay (ML) is the national language, English (EN), Mandarin Chinese (MC), and Tamil are also used in the media. There are more spoken languages (cf. Vollmann & Soon 2018); the ethnic Chinese usually speak South Chinese languages⁡; people will also use Malaysian English ("Manglish", ME) and Bahasa Pasar (BP), creolized forms of the respective lexifier languages. All speakers are multilingual and choose their linguistic resources in dependence of the situation and communication partners.

Among the various ethnic Chinese linguistic groups, the Min (Hokkien, HO), Yue (Cantonese, CT), and Kejia (Hakka, HK) speaking communities are the most numerous in Malaysia (cf. Platt 1977: 365). The speakers of Teochew (TC), Taishan, Hainanese, and others form smaller groups. When different Chinese sub-groups meet, a shared language is needed for communication. This problem has traditionally been solved through multilingual competence, but in recent times, Mandarin is reported to progressively taking this role, even in the family domain, e.g. in Kuching (cf. Ting & Chang 2008).

1.2. Hakka

The Hakkas (客家 Kéjiā), reportedly coming from the North of China, settled in South Jiangxi, Western Min (Fujian), or Eastern Guangdong. Their language is close to the Gan dialects (Sagart 2002) and has converged with Guangdong varieties. From here, many Hakkas migrated to South and South-East Asia, Taiwan, and other places such as Suriname (cf. Leo 2015).

In Peninsular Malaysia, the Hakka speech community is concentrated in Kuala Lumpur, Kajang, Serdang (Seri Kembangan), along the road to Titi Jelebu, Pusing (near Ipoh), also in Bukit Mertajam

---

⁡ These languages or Chinese dialects (方言 fāngyán) are not mutually understandable; nonetheless, they are viewed as being Hán (汉).
and Balik Pulau (Penang). In Sarawak, Hakka is the largest Chinese sub-group (with 160,000 speakers), concentrated in Kuching (Ting & Chang 2008). In Sabah, Hakka even serves as lingua franca (Leo 2015: 76).

1.3. Standard and non-standard languages

In modern societies, it is mandatory for educational and economic success to gain access to an elaborated and reliable linguistic register (cf. Maas 2008, 2010). Usually, a “language” is recognized as such through a process of standardization and propagation of a canonical and elaborated form. All deviation from this canonical form is then called a “dialect”. In a modern society, dialects turn into sociolects representing a smaller identity beside the wider identity represented by the standard language. In traditional societies, the written standard language is distant to all (or most) spoken varieties, used only by few writers (clergy, officials), and therefore does not influence the spoken languages very much; the speakers will simply adapt to communicative needs by learning other varieties, but will usually be excluded from the acrolectal level.

A modern standardized spoken language in combination with school education leads to widespread bilingualism (standard/dialect), access to acrolectal uses and written information, and will gradually take over functions of the spoken varieties, from acrolectal to mesolecetal levels. Spoken varieties will be influenced by the standard language and undergo dialect levelling; their usage will be confined to a basilectal level. In the case of pluricentric languages, alternative standards may develop.

In the case of Chinese, the writing system provides linguistic unity since long. With the definition of modern Spoken and Written Standard Chinese in the 20th century (普通話 pǔtōnghuà ‘the common language’; Mandarin Chinese, MC), a powerful new linguistic resource became available through school education to the speakers of the various Chinese dialects and was readily adopted also in Malaysia. The stratic language situation has influenced the standard language (Malaysian Mandarin, MM), but will also weaken the position of the spoken Chinese varieties.
While languages such as Hakka may be considered *abstand languages*, i.e., varieties too distant to the standard language to be considered a dialect, Chinese identity is inclusive of these speakers.

### 1.4. Chinese identities

For ethnic Chinese, identity is construed as family members (ancestor worship, traditional family structure), as people speaking a ‘dialect’ (e.g., Hakka) and being descendants of a county (e.g., Dàbù, in the province Guangdong/Canton), as Chinese (漢, Hán), and as state citizens of a country (Malaysian) – which is summed up as “Overseas Chinese”. In Malaysia, it is a practical necessity (and prestigious) to speak languages in addition to the one spoken by the group one identifies with, i.e. other Chinese dialects, Mandarin Chinese, and English. The Chinese dialects are neither dialects in the European sense (as a geographical phenomenon) nor sociolects on a social scale (as in urban language research); they are based on the family lineage, ancestry.

The Chinese culture (outside of China) is (often) different from what has been discussed for Western cultures in that ethnic identity (race) is more important than language, religion, or culture (cf. Clammer 1982); there are “ethnic Chinese” people who do not speak Chinese, or who speak a “Chinese dialect” which is not understandable by other Chinese people, but they all adhere to the ethnic group of Chinese, feeling related to Chinese culture.

Identity for Malaysian citizens with Chinese background is therefore layered. Speaking/reading Mandarin refers to their Chinese-ness, speaking/understanding Cantonese is a useful resource (in Kuala Lumpur; but also with (Hongkong) drama series, movies, music, news, written texts). Speaking Hakka (or any other Southern Chinese language) refers to a patrimony/kinship/regional identity. Other languages are acquired if necessary. In general, ethnic Chinese have a pragmatic approach to languages (Platt 1977: 373).

### 1.5. The Chinese writing system

The Chinese writing system is usually only used for Mandarin Chinese, not for, e.g., Hakka (in Malaysia). Due to Hongkong using Cantonese in media, the language users may read articles or news writ-
Multilingualism and language shift

10 in Cantonese. The same Chinese characters can be used for different languages and then be spelled differently; for example a signboard for a shop can be written in Chinese as 福源 (MC fú yuán), but a Hokkien owner will put the English or Malay translation as ‘Hock Guan’, a Cantonese owner will put it as ‘Fook Yun’, whereas a Hakka owner will put it as ‘Fook Yen’. For some lexemes, a different character has to be used. Hakka can be written, too; the following two sentences from our Hakka data are rendered in Chinese characters, with a Mandarin translation (Hakka and Chinese pinyin below):

(01) Writing a non-standard variety (Hakka) in Chinese script
(a) Similar words and writing
    HK: 所以你怎样讲就好复杂啰。
    MC: 所以你说得很复杂了。
    MC: Suò yì ní zhèn me shuō jìu hěn fù zá le.
    gloss: so you such say then very complicated.
    If you say so, thus it becomes very complicated.

(b) Only little similarity (TR101148)
    HK: 个梅菜旧摆孔雀人自家腌个啊，
    MC: 以前他们都是自己腌梅菜。
    MC: Nà xiè méi cài yì qián tā men zi jì yān de a
        gloss: the CLF prsvd. veg. prev.ly 3-PL-people self prsv. REL PTC
        The preserved vegetables that the people made on their own, ...

(c) Free translation of (b) into MC
    HK: 以前他们都是自己腌个啊，
    MC: Yiqian tamen dou shi zijyi yan mei cai,
    before 3-PL all be self pickle prsvd. veg.

In (a), regular phonological input-switch rules (cf. Dressler 1984) allow to write the same word for different spellings; the different words require different written characters. In (b), for instance,个 (MC ga) is used for CLF ge and REL e[55]; 人 (MC rén) represents the same meaning (man/people'), but in different pronunciation (ngin’i) and function (plural marker deu’ngin’i). Finally, only 梅菜 (MC méi cài) is used for the same word ( moi[31] coi[55]). A literal translation is possible with some minor differences (e.g., ge[55] CLF, xié CLF.PL), but the Mandarin translation might better be rendered as in (c).
2. Material & method

There are a number of studies on the use (and possible decline) of the use of Hakka in ethnic Hakka-Chinese families in Malaysia (cf. Platt 1977; Kuang 2002; Ting & Chang 2008; Cheng 2003a,b; Ting 2018). This study describes one extended family (38 individuals, including deceased persons), mainly seated in Kuala Lumpur, and their multilingual linguistic competence in relation to their Hakka identity. The investigation assesses the linguistic repertoire and the linguistic environments of the family members and their linguistic activities in the family and outside, with special attention to the linguistic resources of the younger generations, as well as language attitudes of the speakers.

The first generation (G1) of the Hakka family had migrated from Dābù (大埔 MC Dābù, HK t'ai55 bu44, Guangdong, China) to Kajang (near Kuala Lumpur) at the beginning of the 20th century; the husband later returned to Dābù to build a house for the family members that had remained in China. Generation 2 (G2) was born in Malaya, but one son was sent back to Dābù to repair the family's house and was then unable to return due to the outbreak of WW2. He therefore married a wife in Dābù and returned to Malaya after the war, his wife and two of the three children3 followed shortly afterwards. Generation 3 (G3, 'grandparent generation') were all born in China, but lived in Malaya most of their lives.5 G3 returned to Dābù for a visit in order to meet their relatives.3 Generation 4 (G4, 'parents generation') lives in KL, Penang, and one sister went to live abroad (Taipei, Shanghai). The youngest family members, generation 5 (G5, 'children generation'), are living in KL and in Shāndōng (China). Currently, the family consists of one man and two women from (G3), 5 couples from (G4), and 10 youth and children (G5).

The family members were observed and asked about their language use in 2017 on various occasions in KL and other places, and sometimes via internet communication. Some family members of G3,

3 The visa application took so much time, that another child was born in the meantime; one visa short, the eldest daughter had to stay in China.
4 With the exception of the one daughter in China.
5 HK zon hiong ha 'return to homeland'
G4, and partly G5 provided extensive information both about earlier times and their current language use; some of these conversations have been recorded; not all family members have been asked extensively; instead, few people provided more information about themselves and their relatives. Information about deceased family members were included through reports. Since the assessment of multilingual competence is rather subjective, this analysis gives a qualitative description with examples retrieved during conversations.

3. Analysis

3.1. General description

The oldest family members alive at this point are (G3) one brother and two sisters (2 in KL, one in Dabû). G4 are five sons and two daughters, with their respective wives/husbands. All seven children (G4) were born in KL. G1-G4 are all proficient HK speakers who learned HK as L1, but also learned, to varying degrees, MC at school. G4 are now between 30 and 55 years old.

The family is structured in a traditional way, where the sons (G4) live under one roof with their father (G3). The men (G4) pursue their own businesses or jobs, but also cooperate in many respects. The families of the (G4) women live in modern small families outside of such traditional settings. The father (G3) often stays in a country house and is visited there on weekends by the family members.

In G4, all marriages but one were with non-Hakka persons. The spouse’s mother tongues (G4) are Vietnamese, Iban, CT, and one person has a Taishan/Hakka background “with a Hokkien accent” (i.e., influenced by a HO and TC environment in Bukit Mertajam). The non-Chinese spouses (and some of their relatives) learned MC, not HK. The couples thus speak MC, not HK, with each other.

In G5 (10 persons, aged 4-26), only the elder youth (3 persons) grew up predominantly with Hakka. All other children/youth are considered speakers of MC by the family. The children do not seem to pick up Vietnamese or Iban.

---

6 The smaller children (G5) were not "interviewed".
3.2. Linguistic repertoire of the family members

G1 (great-great-grandparents): The man was a construction contractor in Kuala Lumpur, he spoke Hakka and also at least BP (i.e., not Standard Malay) in order to deal with the local people. The woman was a housewife and spoke only Hakka.

(02) Report from G3 uncle on G1 wife
a24ne31 ma24lai24 vun11du24 m11sid2 q11i ga24 gam31 co24
aunt Malay even NEG know self dare sit
fo24 ca55 hi55 dau55 bin24 sang11
train go DES Penang
Grandma did not even understand Malay, but she dared to take a train to Penang alone.

G2 (great-grandparents) were living and doing business in Setapak where many TC people resided. For business purpose, the man also spoke TC and most probably CT and HO. The woman was a housewife, she could speak HK, CT, and a little bit of BP.

The G3 man (grandfather) learned Chinese and English in school. He took over the business from G2 and lived with Teochew people, was thus able to speak TC. He also knows CT, HO and BP. The G3 woman (grandmother) spoke HK, CT and MM. Although she was helping in the company, she did not pick up HO. Another G3 woman (sister) learned MC and EN at school; she was also helping the brother with the family business and spoke HK, CT and MM. The G3 sister back in Dabù speaks only HK and MC.

G4 (parents) currently consists of eight people in the family house (two died): five brothers, two sisters, one cousin. All people of G4 have learned MC, EN and ML at school, are proficient HK speakers, and all communicate in CT, BP and ME. All men had to travel to different states in Peninsular Malaysia and to Singapore for business, where they picked up HO; two brothers learned some Thai while doing business in the north. The G4 cousin (mother tongue

7 Chinese speakers will often not distinguish between BP and ML, what they are using is rather BP.
8 The women from older generations know fewer languages. For instance, a G3 sister from the mother lineage (not considered here) speaks only HK and little MM.
HK) originally came from Bukit Mertajam where HO is also widely spoken; he also speaks CT. The two G4 women did not pick up HO in spite of working in the family business, because the shop was situated in a HK and CT speaking area. One woman later studied in Taiwan where Min-Nan (Hokkien) is widely spoken; when she then moved to Taiping and Penang where HO is the main language, she picked up the language and is therefore able to understand HO to a certain extent. She does not usually speak this language. While living in Shanghai for five years, she learned to understand Shanghainese. The younger daughter also moved to Penang at a later stage and thereby picked up Hokkien passively.

Among the wives (and ex-wives) of the (G4) brothers, only one is a Hakka woman who knows various languages as mentioned above. The other wives are from different ethnic/linguistic background. One Vietnamese wife learned MM and uses it regularly, the other Vietnamese wife knew EN and subsequently did not acquire much in MC. One wife is an Iban (speaking the Iban language) from a re-mote area in Sarawak. She learned ML and EN at school; Like many Ibans, she learned MM in order to work for Chinese people. She communicates in MM with the family and often uses BP.

The husbands and ex-husbands of the (G4) sisters are also not HK speakers. One speaks CT, MM, ML/BP, and ME; the other husband speaks CT, HO, Teochew, (little/late-learned) MM, ML/BP, ME, and is currently learning Pali (out of religious interest). The third husband is a Hainanese speaker with competence in HK, CT, HO, MM, ML, BP, and ME.

Furthermore, there are two other male Iban speakers in the family's house in Kuala Lumpur; beside Iban, they speak ML/BP, little ME, but also MM.

Among G5 (children). All learn(ed) MC, EN and ML at school. (a) The eldest three G5 were brought up predominantly in a HK-speaking environment both from father's and mother's side. They are working in Kuala Lumpur (CT-speaking area), having no needs to travel to any HO-speaking areas (i.e., they do not speak or understand HO). (b) The two middle-aged (youth) G5 who live in China were always brought up in a Chinese-speaking environment by CT/HK parents not using these languages with them. However, they have obviously picked up HK and CT passively. After learning MC, EN, ML at primary school, they left to reside in China at a young age.
and therefore report not being able to remember ML. In China, they learn MC and EN at school. They understand Shanghainese and picked up Weihai dialect. They can peak in the style of Malaysian English (ME) and enjoy watching Singaporean and Malaysian movies using ME. These children pick up both the linguistic variation in China and in Malaysia. (c) The five young G5 (children) communicate with their parents in MM. The children of the Iban mother may have picked up BP passively, as their mother uses BP very often with her friends.

When summing up the linguistic repertoire of the family members, we get an evenly-spread distribution of linguistic competence (table 03).

(03) Language competence of 38 family members

<table>
<thead>
<tr>
<th>L.g.</th>
<th></th>
<th>lg.</th>
<th></th>
<th>lg.</th>
<th></th>
<th>lg.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HK</td>
<td>24</td>
<td>MM</td>
<td>34</td>
<td>CT</td>
<td>22</td>
<td>HO</td>
<td>13</td>
</tr>
<tr>
<td>TE</td>
<td>2</td>
<td>ML</td>
<td>25</td>
<td>BP</td>
<td>22</td>
<td>EN</td>
<td>24</td>
</tr>
</tbody>
</table>

In spite of being a Hakka family, more family members know Mandarin than Hakka (34:24 of 38), due to non-Hakka family members.

3.3. Multilingual practices

The enumeration of many languages spoken by almost all family members hides the fact that these languages are not all known to the same extent. What certainly helps learning languages in Malaysia is the high degree of grammatical convergence of these languages, including South Chinese languages, the Malaysian form of Mandarin (MM), Malaysian English (ME), and Bahasa Pasar (BP). Language learning may predominantly mean phonological and lexical adaptation.

(04) Word-by-word translatability of languages (HK, ME, MM)

(a) Speaker LSM (G4) and translations

HK: ai25 ngid2 nga24 pen11 iu24 mai24 loi11 bun24 ngai11 e55.

ME: That day my friend buy and give me one.

MM: nà tiān wǒ de péngyǒu mài lái gěi wǒ de.

9 In this count, the children from 6:00 onwards have been considered knowledgeable of English, as they learn English at school; of course, the younger ones do not yet know much.
Many sentences are structurally translatable word-by-word (a,b – HK/MM/ME), some seem to be more different (c – HK/ME); but we can see the indeclinability of the English verbs (give, is) and also their orientational neutrality (aunt give = aunt was given), with ‘already’ as a functional word equivalent to perfective \( T \) (MC le; HK liaohoi). The speakers also follow a Chinese topic-comment structure in ME. As mentioned before, Mandarin has developed into a Malaysian variant which is recognisable by other Chinese speakers, through spellings, lexical items, loanwords, and style/structure.

The constant presence of various languages leads to conversations that are often bi- or even multilingual with frequent code-switching. The conversations also reveal that the children (G5) hear Hakka words.

(05) G4 to G3, loanwords in Hakka

\begin{verbatim}
ba\(^4\), ten\(^2\) gong\(^3\) ngid\(^2\) ngai\(^1\) dai\(^5\) ngi\(^1\) hi\(^5\) Bukit Bintang
dad, tomorrow I take you go Bukit Bintang

HK HK HK HK HK ML

sid\(^2\) bufei.

eat buffet

Dad, tomorrow I take you to Bukit Bintang to have buffet.

HK EN

(06) G4 to G5, Mandarin, with kinship terms mostly in Hakka

\begin{verbatim}
Qú jiào ní de fèi bag\(^2\) bag\(^2\), da\(^2\) da\(^2\), xiào shūshú hé
go call your fat fa.eld.bro, grandfa. small fa.yo.bro and
MC MC MC MC HK MC HK MC

sug\(^2\) me\(^2\) lái chi satay.

fa.yo.bro.wife come eat satay
HK MC MC ML

Go call your fat uncle (father’s elder brother), grandpa, youngest
uncle (father’s younger brother) and aunt (father’s younger brother’s
wife) to come to eat satay.
\end{verbatim}
The competence in other Southern Chinese languages varies, with TC and HO being least known, but a good command of CT by many family members. The men need to use other South Chinese languages for business purposes; the women by and large report not understanding other dialects such as HO; but they also seem to know “something”, be it only from the fact that HO is an important influence on Bahasa Pasar which they can speak. In short, these languages somehow converge into a general communicative competence where word or language choice are basically pragmatically determined. Languages are picked up when necessary, and can also be forgotten, when no longer needed. Speaker (G3) reports he used to speak TC, but when he tries now, he ends up speaking HO:

(08) Speaker (G3) on his competence

m\textsuperscript{11} rh\textsuperscript{4}ng\textsuperscript{5}e\textsuperscript{55}, rh\textsuperscript{4}vui\textsuperscript{55} ng\textsuperscript{11}e\textsuperscript{55} shi\textsuperscript{11}gien\textsuperscript{24} bien\textsuperscript{24}ma\textsuperscript{2} not same-REL, because you-REL time change-PTC.

rhi\textsuperscript{11}qian\textsuperscript{11} nga\textsuperscript{4}deu\textsuperscript{1}rh\textsuperscript{11} du\textsuperscript{4} m\textsuperscript{11} ... rh\textsuperscript{11}ga\textsuperscript{24} nga\textsuperscript{4}deu\textsuperscript{24} in the past I-PL-people all not, people I-PL.

tung\textsuperscript{11} chau\textsuperscript{11} zeu\textsuperscript{24} rh\textsuperscript{11} cu\textsuperscript{55}, cu\textsuperscript{55} hoi\textsuperscript{24} gid\textsuperscript{7} do\textsuperscript{24}ng\textsuperscript{11} in with Chao.Zhou people live, live already so many year, gong\textsuperscript{31} chau\textsuperscript{11}zeu\textsuperscript{24}fa\textsuperscript{55}e\textsuperscript{55} jiu\textsuperscript{15} rh\textsuperscript{11}sien\textsuperscript{55} go\textsuperscript{55} gin\textsuperscript{2}nga\textsuperscript{55} gong\textsuperscript{31} talk Teochew-lg.-PTC just fluent. now speak chau\textsuperscript{11}zeu\textsuperscript{24}fa\textsuperscript{55} du\textsuperscript{24} gong\textsuperscript{31}dau\textsuperscript{31} fug\textsuperscript{2}gien\textsuperscript{55}fa\textsuperscript{55} m\textsuperscript{11} hiau\textsuperscript{31} Teochew-lg. all speak-DST Fujian-lg. not can gong\textsuperscript{31} chau\textsuperscript{11}zeu\textsuperscript{24}fa\textsuperscript{55}.
speak Teochew-language.

(100) **G4 to G4, code-switching (CT, HK, HO)**

I am sorry PTC I today not free drive you

CT CT HK HK HK HK HK
lu chiah ka-ki la.
you eat self PTC

FJ FJ FJ FJ

I am sorry. Today I have no time to give you a lift. You take care of your own (transport).
3.4. Language choice

The standard Malay language, although learned at school, seems to be used only in rare, specific circumstances (e.g., in state offices); even then it may often be used in the form of BP rather than ML; the dialectal/spoken ML is not always understood (in other places, such as Penang, Kelantan, Terengganu).

While English is a good candidate for inter-ethnic communication, it was not used very much during the observation, except for loanwords. English was understood, but people were often shy to use EN with a non-Malaysian person.

Inside the family, the Hakka family members (G4, G3) speak HK with one another, and MM with the others (spouses, G5). The non-Hakka family members and the youngest children do not acquire an active use of HK, and speak MM in the family.

Outside the family, CT and MM are most relevant. CT is also used with family members of CT background (in-laws), with friends, and at the workplace. HK or CT are used in the market with sellers known to the speakers. The initial choice of language with unknown people seems to be delicate; speakers will initiate the conversation with a Malaysian form of a standard language, not the educated variant: MC, ME, or BP. If there are indications of the other person being, e.g., a CT or HK speaker, the language may be switched. The choice of a common language serves to create greater intimacy.

While it is not yet clear whether the young children will also start to actively use HK or not, there is a difference inside G5 between the elder children (18-26) being HK speakers and the younger children (4-17) being speakers of MC.

[...], like my own children, ehm, the father is Cantonese speaker, and I am Hakka speaker, but I decided to speak only Mandarin with them since they were small. But they heard me talking to my father and mother in Hakka. So, surprisingly now I found out that they were able to speak Hakka. (Mother, G4)

Among the younger children (G5), all but one are girls; It was observed that the grandfather (G3) and father/uncles (G4) do speak Hakka to the youngest boy (G5, 4 years), but not to the girls. This may be attributable to the concept of the male family lineage.
3.5. Attitudes and identity

When asked about their attitudes and identities, the (G3/G4) adults present as Malaysians -- which is maintained when living abroad, e.g., in China. This identity is not reflected in the use of the Malay language. While the Hakka-speaking family members (G3, G4) have a positive attitude towards Hakka as an identity marker for the family when asked, this is usually not discussed in daily life; Hakka is not considered a “language” as compared to Cantonese or Mandarin.

Hakka identity is a matter of clan/ancestor affair (宗族 zōng zú) for the male lineage with roots in Dābū, the Hakka language is a symbol, but not constitutive, of the family identity. Chineseness is the ethnic and cultural identity (中国人 huá rén), linguistically represented by the Chinese written language.

G4 speakers expressed the view that they do not speak a ‘real’ Hakka, but just a Malaysian, modified version. Due to a visit to Dābū, some of the family members are aware of important linguistic differences between themselves and their relatives in China. This is common for migrants who do no longer participate in the linguistic dynamics of their homeland, but adapt to their new environment instead. Original Hakka dialects converge into a Malaysian Hakka.

(09) G4 speaker on his/her competences in Hakka

I grew up speaking only Hakka. I learned Mandarin when I went to school at the age of 7; I often went staying with different relatives during holidays. I was very much exposed to the Hakka spoken by grandmothers, grandaunts, uncles and aunts. I learned Cantonese from school mates and TV programs. I remembered watching a TV show called ‘Empat Sekawan’¹¹. There were four characters in this show, each of them spoke one different dialect; e.g., Lai Meng spoke Hakka, Hon Ying spoke Cantonese, Wong Ho spoke Gaozhou, Hai Yong spoke Hokkien. I could understand Hakka and Cantonese easily. As I had limited vocabulary in Hakka, whenever I needed to

---

¹⁰ The family looks back at an ancestry line of 21 generations, the people are known by name since the 16th generation. The family has a genealogical chart after which the sons of each generation are named. The respective fathers followed this custom for boys, and did not name the girls in the family with these names.

¹¹ ‘Empat Sekawan’, also known as ‘Sei Hei Lam Mun’, this sitcom by RTM was one of the most popular shows in the 60’s.
explain something in Hakka and lack of that word, I would replace
with Mandarin or Cantonese. That’s why I think my Hakka is not as
original as my parents’, uncles’ and aunties’.

G3 speakers are also aware that their language has changed.

(10) G3 on whether he can speak like the people in Dābù (TR201175-7)
gi55 ded2 loi11 jiu55 he25 gong35 ta13 bu44 fa55 lo55. gin24 nga55
remember come then be talk Dabu lg. PTC now
jiu55 he55 m11 gi55 ded2 ma24. ngiong31 men24 hian1 gi55 ded2?
just be not remember PTC how can remember
(If I) would remember. I would talk in Dabu language. Now (I) just
cannot remember. How could (I) remember?

Hakka cannot be read or written by the family. It has no other merit
than giving a familial identity. Due to its weak basilectal status, the
parental desire for education and social advancement leads to favor-
ing Standard Mandarin (for some even Pǔtōnghuá instead of Malay-
sian Mandarin) in child upbringing (disregarding the sociopragmatic
functions of multilingual code-switching).

(11) G4 speaker on child-directed language choice
[...] and I want my children to speak "pure" Chinese which means
the "standard", because usually people in Malaysia speak with a
mixture of other dialects, and also not the right tone. Yah, so, I
would prefer my children to speak standard Mandarin. That’s why I
started with Mandarin, only Mandarin. (Mother, G4)

Like in other modern societies, educated urban families will tend to-
wards a standard language which may eventually lead to language
shift and a remodelling of the sociolectal situation (giving up a mini-
ortality language and acquire sociolects of the new language). In the case
of “Mandarin”, a new variation between foreign standards (China,
Taiwan) and a domestic standard (Malaysia) with dialects and socio-
lects (“Malaysian Mandarin”, influenced by different substrate lan-
guages) replaces the functions of the traditional linguistic variation.

(12) G4 speaker on recognizing dialectal differences
Sometimes you can tell from their Mandarin, sometimes you can
tell from their Chinese that this person is speaking, you know,
roughly which area is she from. From the northern part or the sou-
thern part, or she is very much influenced by Cantonese – you can
tell that.
Speaking Standard Mandarin is not well-received among Malaysian Chinese in most situations. Highly educated speakers of Mandarin need to attune to the Malaysian standards in everyday life, or else they may be perceived as foreigners or as snobbish/pretentious. On the other hand, educated speakers may attune to Pǔtōnghuà or Taiwanese Mandarin, when they are meeting with the respective groups. Taiwanese and Chinese people may sometimes have troubles understanding a speaker of Malaysian Mandarin (MM).

3.6. Language change in progress

Since all Chinese languages do not diverge very much in grammatical constructions, convergence is very easy in the contact situation. For instance, the HK particle liao in G3 is often replaced by hoi in G4 which seems to be a (Malaysian) CT influence (for MC 去 le); the grammatical categories expressed by liao/le/hoi are identical in all three languages.

(13) G3 and G4 in free conversation; usage of liao and hoi;
(a) sai⁵⁵ gön⁴⁴ liau⁴⁵ a⁵⁵ ... speaker G3
    sun-dry dry LE PTC
    After having dried under the sun, ...
(b) gin⁴⁴ nga⁵⁵ gia⁴⁴ lau⁴⁴ gung⁴⁴ go⁵⁵ sin⁴⁴ hoi⁴⁴, ... speaker G4
    now 3PSS old-male pass.body LE
    Now her husband has passed away, ...

Dialect levelling takes place by replacing HK words with MC words that are phonologically integrated into HK. The following example shows the replacement of the traditional expressions by a MC loan-word:

(14) Dialect levelling in Malaysian Hakka (G3 = a,b and G4)
(a) ngai¹¹ oï⁵⁵ hi⁵⁵ si¹¹ hang⁴⁴. [= 屎坑 shìkēng]¹²
    I need go shit drain.
(b) ngai¹¹ oï⁵⁵ hi⁵⁵ siao³¹ pen⁴⁴ fong⁴⁴. [= 小便房 xiǎojiān fāng]
    I need go small.business room
(c) ngai¹¹ oï⁵⁵ hi⁵⁵ ci⁵⁵ so³¹. ← 厕所 cè suō
    I need go toilet
    I need to go to the toilet.

¹² A kind of earth closet; but also the squatting toilet could be named such in earlier times. G2 (from China) is remembered having used the word pun⁵⁵ gong⁴⁴ (Chin. fèn gang).
The word *siao*³⁷ *pen*²⁴ *fong*²⁴ ('small business room') would be a good word for toilet (for G3 speakers); the word *si*¹¹ *hang*²⁴ ('shit-drain') is no longer usable, as it refers to the traditional kind of toilet. G4, however, prefer the word *ci*⁵⁹ *so*³¹ borrowed from MC *cè suò* ('toilet') and adapted phonologically (loanword integration).¹³

MC words can easily be integrated into HK due to regular phonological and tonal correspondences (input-switches). The fact that MC has a fully developed lexicon for modern things and has prestige makes it a useful resource. In this way, smaller languages may converge and transform into mere sociolects, varieties which are systematically dependent on a standard.

Differences between the HK of G3 and G4 can be observed in abundance; for instance, G3 has variants of some words where G4 has only one form; loanword integration (from non-Chinese languages) is stronger in G3 speakers (e.g. *[sun]³⁴*ma²³*] ← ML *[souma]* 'semua'; *[ba²⁴sag²⁴]* ← ML 'pasar'), while G4 uses EN and MC words in an adapted HK phonology (e.g. *[bebi]* 'baby'; *[plen]* ← 'plan'). Code-switching using EN is not seen among the older speakers (G3), but happens with younger speakers (G4).

(15) Loanword integration in G3 (a), and borrowing in G4 (b)

(a) aï you, san²⁴ hong³⁵ e³¹ sui³¹ sung²⁴ ma²⁴ leu⁵⁵ log⁵⁵
   ITJ mountain on REL water all flow down
   there PTC.
   The water from the mountain all flows down there.

(b) gia²⁴ ji³¹ gien³⁵ go³⁵ gi¹¹, in²⁴ vui³⁵ gi¹¹ he³¹ bebi ma²⁴
   3 eld.sis see PFT 3 because 3 be baby PTC
   Her elder sister had seen her, because she was a baby (at that time).

Both G3 and G4 use the many particles Malaysian languages are known for. The syntax is not different between G3 and G4, but the lexemes sometimes are.

Phonological differences between G3 and G4 are also observed; dialect levelling seems to occur in the direction of MC; for instance, the HK monophthongs which are diphthongs in MC get changed

---

¹³ This example can be compared to, e.g., Austrian *[abu:t]* ('Abort' ← lat.) which is no longer used since two generations and was replaced by *[kl:o:]* (← 'Klosett' ← lat.) which is again replaced by standard German *[tœ'lete]* ('Toilette' ← fr.).
from G3 to G4: lo⁵⁵ fu⁴¹ → lau⁵⁵ fu⁴¹ (tiger; cf. Md. 老虎 lăohù), lo⁵⁵ zu⁴¹ → lau⁵⁵ su⁴¹ (mouse, cf. Md. 老鼠 lăoshù). More observations:

(16) Phonological differences observed in the corpus

<table>
<thead>
<tr>
<th>G3</th>
<th>G4</th>
<th>TRL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ki⁵⁵</td>
<td>hi⁵⁵</td>
<td>go</td>
</tr>
<tr>
<td>mag⁶⁵ pien⁴¹</td>
<td>mag⁶⁵ pen⁴¹</td>
<td>cereal</td>
</tr>
<tr>
<td>ngiong⁶⁵ ben⁴¹</td>
<td>iam⁶⁵ men⁴¹</td>
<td>how</td>
</tr>
<tr>
<td>ngien¹¹</td>
<td>nen¹¹</td>
<td>year</td>
</tr>
<tr>
<td>m¹¹ mai⁵⁵</td>
<td>m¹¹ moi⁵⁵</td>
<td>don't want</td>
</tr>
<tr>
<td>do⁵⁵</td>
<td>dau⁵⁵</td>
<td>arrive</td>
</tr>
<tr>
<td>mo¹¹</td>
<td>mau¹¹</td>
<td>not have</td>
</tr>
<tr>
<td>gin⁴⁴ nga¹¹</td>
<td>gin⁴⁴ nga⁵⁵</td>
<td>now</td>
</tr>
</tbody>
</table>

3.7. Borrowing

One important difference for Malaysian Hakka vs. other variants in mainland China or Taiwan come from the integration of EN and ML loanwords. Again, many such words have to do with the authorities, such as here:

(17) ki bun mada lokap hoi. He has been detained by the police.

In this sentence (G3), the word [mada] is from (spoken) ML mata-mata (‘police’), [lokap] obviously is from EN ‘lock up’. The HK verb [bun] ‘give’ as a PSV marker corresponds with MM 给 géi / MC bèi (被), [hoi] equals MC le (了).

Other examples (in G3) from this sphere are [məd-ʣə-la] from ML bicara (‘negotiate in court’) and [lɔ] (law) itself; G4 uses HK [hi goŋ haj] (‘go talk’) which copies MC qù tān (去谈), or MC qù shuō shuō kān (去说说). The ML words salah ‘wrong, unlawful’, suka ‘like’ (ML sukɔ → HK su⁴⁴ gag⁴¹), pandai → HK ban¹¹ nai⁵⁵ ‘clever, good at’ are used in HK:

(18) ML salah, suka, pandai in HK
(a) mau⁴¹ dai⁵⁵ IC cud⁴⁴ mun⁴⁴ salah e⁵⁵ a⁵⁵.
VNEG bring IC out door wrong REL PTC
It is unlawful to go out without your identity card (IC).

G4 may also use [pak’ jiuf] (traffic cop), in other languages, it will be [polis] or jingchá 警察 (MC) instead of [mada].
The car sphere has many loanwords: roun(d)abou(t), [streɪŋ] (steering), [brɛk] (brakes), [ovəhɔ] (overhaul), [taɪn bɔntʃet′] (tire puncture), [loli] (lorry), [ɔt′] (road; also ML jalan in G4, eventually HK lu (ㄌ ), [gəladzi] (clutch), [da signər] (HK give + signal), [bomba] (bumper), [esəba] (shock absorber), [ɛndzin] (engine), [han haiwei] (go by the motorway), [go tɔ] (pass the toll), [teksi] (taxi), modobai (motorbike), [jenbai] (exhaust pipe, MC yān 煙 'smoke' + pipe), [laison] (license\(^\text{\textsuperscript{15}}\)), [bomi?] (permit), [tʃatʃ]\(^\text{\textsuperscript{16}}\) ten (charge + MC battery), [ka:pak′] or [pakin] ('(car) park(ing)'), [bun pakin lue] (HK 'give' + EN 'parking' + HO/HK 'money').

Furthermore, many other EN words found their way into Hakka: [aisgəlim] (ice-cream), [missi] (nurse, from miss), [bɛbi] (baby), [aβai] (nickname for small boy), [aʃet] (nickname for a fat person), [kɛk] (cake), [cukulu] (chocolate), [kɔpi ɔ/kəɔ] (black/milk coffee; EN+HO), park, [let′ tai] (neck-tie), [nga lak bui] (glass+ 杯), sofa, [id dzag bia tɕui] (one juq\(^\text{\textsuperscript{17}}\) of beer),\(^\text{\textsuperscript{18}}\) [babik′u] (BBQ), [tsoʔ ti-sat] (wear t-shirt), [(na) səbana] (take the) spanner; from ML: [(zɔn\(^\text{\textsuperscript{17}}\)) gambong] (return to) home/town; Mal. kampung 'village/ hometown'), [hi ba:sag′] (go to pasar), [ba:sag′ malam] (night-market).

4. Discussion

In general, modern spoken standard languages quickly gain ground at the expense of smaller languages and dialects everywhere. School education, parental aspirations, and media consumption are all in favor of the standard language. Spoken varieties are then subject to dialect levelling and eventually their decline (language shift).

\(^{15}\) Also used for passport, etc.

\(^{16}\) Phonological loanword integration depends on the degree of acquaintance with English; therefore [tʃat′] or [tʃa(ː)ʃ].

\(^{17}\) Not the HK classifier zəg, but EN 'jug' (as a measure word).

\(^{18}\) cf. MC 啤酒 ｐi jiú 'beer' (lit. beer+alcohol).
The Hakka language has experienced changes in the migratory situation, adapted to the multilingual environment (convergence, borrowing), and converges with the standard language which in turn has adapted to the multilingual setting of the local environment (MM).

We see a three-generation decline of Hakka competence from fully proficient grandparents to still proficient, but insecure adult speakers, to young children who are no longer expected to speak the language and do not actively use it. This development is favoured by interethnic marriages and the educational offer of a standard language which serves needs of wider communication better and gives access to media, education and economic opportunities.

5. Abbreviations

3 pronoun 3rd person LE Chinese ҉ le particle
BP Bahasa Pasar lg. language
Chin. Chinese LSM a speaker ID
CLF classifier ML Malay(sian) language
COS change-of-state 了 MBW a speaker ID
CT from Cantonese MD Mandarin
DST destinative particle ME Malaysian English
EN English MM Malaysian Mandarin
fr. French PL plural particle
HO Hokkien (Fujian) PTC (discourse) particle
G1, etc. generation 1, etc. REL connection particle
HK Hakka SRC source
IC identity card Std. standard
ITJ interjection TC Teochew
KL Kuala Lumpur TRL translation tier
LE gramm. particle 了 YPK a speaker ID
lat. Latin

6. References


